1212–1219 Nucleic Acids Research, 1996, Vol. 24, No. 7

Sequences attaching loops of nuclear and mitochondrial DNA to underlying structures in human cells: the role of transcription units

Dean A. Jackson, Jon Bartlett and Peter R. Cook*

CRC Nuclear Structure and Function Research Group, Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

Received January 12, 1996; Revised and Accepted February 15, 1996

GenBank accession numbers+

ABSTRACT

DNA sequences attaching loops of nuclear and mitochondrial DNA to underlying structures in HeLa cells have been cloned and 106 representative clones sequenced; 10 clones containing random genomic fragments served as controls. As chromatin is prone to rearrangement, care was taken to isolate sequences using 'physiological' conditions that did not create additional attachments. Comparison (by Southern blotting) of the concentration of each cloned sequence in 'total' and 'attached' fractions of DNA showed that most clones did contain attached sequences, but even highly-attached sequences were not attached in all cells in the population. Results demonstrated that 28% of clones were derived from three specific parts of the mitochondrial genome and 22% from different parts of the *alu* repeat. In addition, 41% of clones contained unique nuclear sequences; these contained no more of the motifs found attached to nuclear scaffolds or matrices (ie SARs or MARs) than would be expected from their base composition. No other attachment motif(s) could be identified by sequence analysis. However, Northern blotting showed that all the mitochondrial clones and 76% of clones containing unique sequences were transcribed; the degree of attachment correlated with transcriptional activity. These results are consistent with transcription being responsible for ever-changing attachments in both nuclei and mitochondria.

INTRODUCTION

It is widely assumed that nuclear DNA is looped into domains by attachment to some underlying skeleton (e.g. a matrix or scaffold; for reviews see 1–3). Most models for loop structure involve stable 'structural' attachments of repeated DNA motifs in most cells in an organism, as well as 'functional' attachments that vary from cell to cell, depending upon replicational and transcriptional activity (4).

Three general approaches have been used to define attached sequences. One involves progressively detaching DNA from the skeleton with a nuclease; sequences at attachment points resist detachment and become enriched in a pelleted fraction (5,6). Sequences with a high affinity for isolated skeletons can also be selected by incubating isolates with various DNA fragments to see which bind specifically (7). And if attached sequences determine the functional boundaries of a chromatin domain, they can also be defined by their ability to buffer the domain from the repressive effects of flanking chromatin (8–10).

Despite the availability of these assays, there remains little agreement on the identity of the molecules at attachment points. Sometimes the DNA sequences that resist detachment are found within genes (e.g. the canonical matrix attached region, or MAR, of the kappa immunoglobulin gene; 7), sometimes not (e.g. the canonical scaffold attachment region, or SAR, next to the histone gene; 6); sometimes MARs/SARs act as domain boundaries (e.g. the chicken lysozyme MAR; 9), sometimes not (e.g. the MAR/SAR of the heat-shock locus; 10). Moreover, different proteins with a high affinity for MARs/SARs have been identified, including topoisomerase II (11), SATB1 (12), SAF-A (13) and ARBP (14). In the absence of any consensus, sceptics suggest that these differences result from the induction by the unphysiological conditions used during isolation of new and artifactual attachments of loops to an underlying substructure (2,15).

Unphysiological conditions are often used because nuclei and chromatin tend to aggregate at an isotonic salt concentration. However, aggregation can be prevented by encapsulating cells in agarose microbeads (diameter 50–150 μm) before lysis with Triton X-100 in a 'physiological' buffer (16); then the now-encapsulated chromatin is accessible to molecular probes like enzymes and antibodies. The loops are sufficiently protected to retain their attachments, integrity and contour length during extensive manipulation (17). We have now cloned and sequenced DNA fragments at attachment points in this material. We find no evidence for 'structural' attachments involving the same (repeated) motif; rather, different sequences had different probabilities of attachment,

* To whom correspondence should be addressed

+X89762–X89767, X89831–X89840, X89842–X89848, X91562, X91844, X91537–X91561, X91563–X91585, X91587–X91610, X91841–X91844

a high probability correlating with a high transcriptional activity. Therefore attachments are predominantly 'functional'.

MATERIALS AND METHODS

Preparation of the library

Unsynchronized HeLa cells were grown in suspension for 24 h in [^{3}H]thymidine (0.1 μCi/ml; ~50 Ci/mmol) to label uniformly their DNA; this allows the percent chromatin remaining attached during isolation to be estimated from the percent acid-insoluble [^{3}H] remaining. Labelled cells were encapsulated in agarose (2.5 × 10^6 cells/ml packed beads), washed in a 'physiological' buffer (PB; final concentrations are 22 mM Na^+, 130 mM K^+, 1 mM Mg^{2+}, <0.3 μM free Ca^{2+}, 132 mM Cl^-, 11 mM phosphate, 1 mM ATP and 1 mM dithiothreitol; pH 7.4), permeabilized by washing 3× in PB + 0.25% Triton X-100 and then rewashed 5× in PB. Beads, 2 ml (containing 5 × 10^6 cells), were resuspended in 10 ml PB, incubated (20 min; 33°C) with *Hae*III (5000 U) and *Eco*RI (25 000 U), washed in PB, reincubated with the two enzymes as before, and released chromatin fragments electroeluted (3 V/cm; 5 h). Beads, 2 ml (now containing 4.3% chromatin), were incubated in 10 ml PB with *Alu*I, *Hin*fI, *Rsa*I, *Pst*I (500 U each), *Hpa*I, *Msp*I, *Mbo*I and *Sau*3A1 (250 U each), placed in dialysis tubing, re-subjected to electrophoresis (2 V/cm; 4 h) and both beads and surrounding fluid recovered. DNA was purified by treating beads with 0.2% sarkosyl, RNase A (50 μg/ml; 33°C; 30 min) and proteinase K (200 μg/ml; 33°C; 1 h); after melting beads (70°C; 15 min), the solution was extracted successively with phenol (this removes agarose), phenol–chloroform and then chloroform and DNA precipitated with ethanol. The ends of fragments (number-average size 1.2 kb) were 'in-filled' using 'Klenow', ligated into Bluescript II vector (cut with *Sma*I and phosphatase-treated) and used to transform *Escherichia coli* strain DH5 to give ~5000 colonies (18; inserts had a number-average size of 1.0 kb). A similar library was also prepared from 'total' DNA (Fig. 1, stage D).

For Figure 1I, DNA was purified from (i) beads collected immediately after the second incubation with *Hae*III and *Eco*RI (total DNA, *Hae*III–*Eco*RI fragments; lane 1); (ii) beads immediately after the first electroelution (4.3% of *Hae*III–*Eco*RI fragments remaining attached; lane 2); (iii) buffer outside beads in the dialysis bag after the second electroelution (2.9% released fragments; lane 3); and (iv) beads in the dialysis bag (1.4% attached fragments; lane 4). The number-average molecular weights of fragments and corresponding loop sizes were determined (17,19) using agarose gels (1.7% in TEA buffer; 30 V per 10 cm, 5 h).

Sequencing

Sequences were determined after amplifying inserts in clones using the polymerase chain reaction and Bluescript primers, using 'Sequenase' (USB) and 6% polyacrylamide gels (18). Clones, 140, containing 'attached' fragments (designated by C#) were amplified; 106, 27 and 5 had inserts of <500, 500–2000 and >2000 bp, respectively (2 had no inserts; number-average size of inserts was 0.5 kb). The 106 with short inserts were sequenced from both ends; most inserts were short enough to allow complete overlap of the two resulting sequences (designated as LAS#), but longer ones gave two partially or completely non-overlapping sequences (designated as LAS#A or B) that were analysed independently. Sixteen inserts (8, 3 and 5 homologous to mitochondrial, *alu* and other sequences, respectively) were identical to others in the 106 clones and so could have arisen by division of bacteria following the same initial ligation event; they were not analysed subsequently. The resulting loop attachment sequences (LASs) were analysed using the GCG package (program manual for the Wisconsin GCG package, version 8) for homology with sequences in the EMBL database version 44.25 were identical (or differed at 1–2 places) with mitochondrial sequences. Twenty, 3 and 3 clones were homologous with *alu*, alphoid and LINE repeats, respectively. The TIGR database of ESTs (20) was then screened with the non-mitochondrial and non-repeated sequences, including unique sequences 5′ or 3′ to regions of homology with *alu*-a (named LAS#END1 and END2, respectively). Sequences were also analysed using the GCG and Staden (21,22) suites, as well as PROMOTER SCAN (23), and deposited in the EMBL database with accession numbers X89762–7, X89831–40, X89842–8, X91562, X91844 (mitochondrial), X91537–61, X91563–85, X91587–610, X91841–4 (non-mitochondrial).

Mapping proximity to attachment points

The relative concentrations of each LAS in 'total' and 'attached' DNA fractions were determined by 'Southern' blotting. Attached DNA was isolated as illustrated in Figure 1E after treating (20 min) beads (5 × 10^6 cells/ml beads) with *Hae*III (250 U/ml) and *Eco*RI (2500 U/ml). This modification was necessary as exhaustive cutting as in Figure 1E is expensive; as a result, ~90% chromatin is detached. Attached DNA was then purified as above, as was total DNA after adding Sarkosyl to unencapsulated cells. 'Total' and 'attached' fragments were completely recut with *Hae*III, precipitated, appropriate amounts of ^{3}H loaded on a gel, subjected to gel electrophoresis (all as above), blotted, hybridized, autoradiograms prepared using X-ray film or a PhosphorImager, and enrichments of 'attached' bands relative to those in 'total' DNA determined by quantitative densitometry (24). ^{32}P-Labelled probes (specific activity ~10^9 c.p.m./μg) were prepared using a 'random primed DNA labelling kit' (Boehringer). Note that depletions cannot be quantitated accurately, as relevant concentrations of 'total' DNA were not run in adjacent tracks. Therefore all depletions (and enrichments of <2×) are (conservatively) grouped together; this has the effect of underemphasizing the number of LASs that are attached.

The relative concentrations of sequences in the 'total' clones was determined in a way that overemphasised their attachment. Many of the *Hae*III fragments contained in the 'total' library are <300 bp and so are lost during blotting. Therefore larger 'total' and 'attached' fragments were prepared as in Figure 1D and E using only *Eco*RI; ~20% chromatin was then retained and this was used as the 'attached' fraction for blotting. Then, a greater proportion of the sequences in the genome appear enriched.

Measurement of transcript concentration

The concentration of transcripts in whole cells, nuclei and polyA^+ RNA that were complementary to LASs was determined using 'Northern' blots (18). Cells were washed in PBS, and nuclei isolated by swelling cells (15 min, 4°C) in 50 mM NaCl, 1 mM $MgCl_2$, 10 mM Tris (pH 8.0), breaking them with 20 strokes of a Dounce homogenizer, before spinning (2000 *g*, 5 min) and rewashing. RNA was extracted from cell or nuclear pellets using a RNAzol B kit (Biogenesis); polyA^+ RNA was selected from

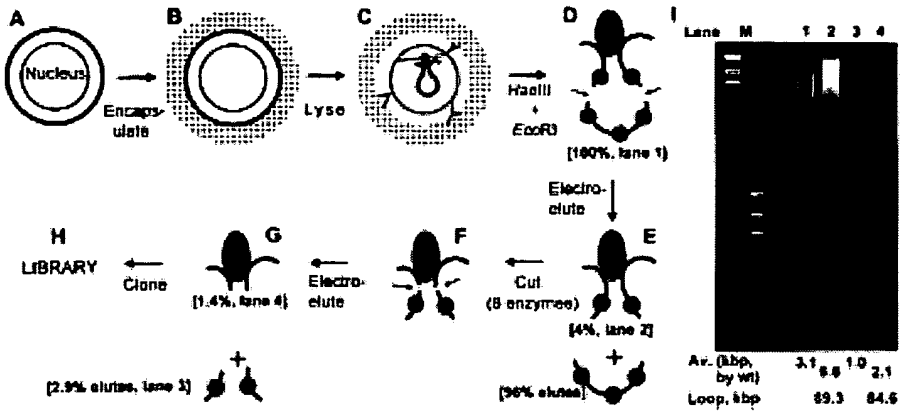

Figure 1. The approach used to create a library containing sequences at attachment points. The percentage of chromatin present at various stages, and the origin of the samples applied to the different lanes in (**I**), are indicated in square brackets. (**A**) Unsynchronized HeLa cells were grown (24 h) in [^{3}H]thymidine to label uniformly their DNA; the percentage of chromatin remaining at different stages in the procedure could then be estimated from the amount of acid-insoluble ^{3}H. (**B**) Cells were encapsulated in agarose beads. (**C**) After lysis with Triton X-100, the cytoskeleton, lamina, internal nucleoskeleton, associated attachment points (oval) and DNA loops covered with nucleosomes (small circles) became accessible to added enzymes; mitochondrial genomes (not shown) are also attached to cytoskeletal elements. (**D**) Added *Eco*RI and *Hae*III cut (arrows) chromatin loops (shown enlarged); after pelleting, this cutting was repeated. (**E**) 96% chromatin was eluted from beads to leave 4% attached. (**F**) Residual fragments were trimmed (arrows) closer to attachment points with eight other restriction enzymes. (**G**) Elution removes 2.9% chromatin to leave 1.4% attached. (**H**) Residual fragments derived from both nuclear and mitochondrial loops were cloned. (**I**) Samples (0.5 µg DNA; 7500 c.p.m.) from different stages in the procedure (plus λ/*Hin*dIII and φX174/*Hae*III markers, M) were subjected to electrophoresis on a 1.7% agarose gel, stained with ethidium and the gel photographed under uv-illumination. Weight-average molecular weights and loop sizes are indicated.

whole cell RNA passing twice over oligodT-cellulose (Pharmacia). RNA samples were precipitated, redissolved in formaldehyde gel-loading buffer and run (2 V/cm; 18 h) in 2% agarose-formaldehyde gels. Gels were either washed, stained with ethidium and photographed (e.g. Fig. 5, left) or RNA was transferred to nitrocellulose (0.45 µm pores; Schleicher and Schuell) and hybridized with probes made as above; finally autoradiographs were prepared. For Figure 6, the strongest signal—whether in whole cell, nuclear, or polyA$^+$ RNA—was expressed on the scale reflecting the minimum exposure required to detect a band on film (PhosphorImager exposures used for weak signals were converted to an equivalent time for film): – (no signal), + (>5 d), ++ (16 h–3 d) and +++ (2 h).

RESULTS

The approach

Figure 1 illustrates the approach used to create a library containing sequences at points of attachment. HeLa cells were encapsulated in agarose, lysed, and chromatin loops cut exhaustively with 10 different restriction enzymes; then most chromatin was electroeluted from the agarose beads to leave only 1.4% of the original amount. Finally, DNA was purified from these residual fragments and cloned. As mitochondrial remnants remain associated with the cytoskeleton and resist elution (25), sequences involved in attaching the mitochondrial genome to the substructure are also isolated.

This approach requires that existing attachments are not broken, nor new attachments created, during the procedure; then, the contour length of the loops should remain unchanged. Therefore we monitored the length of nuclear loops, which can be calculated from the percentage of chromatin (i.e. [^{3}H]DNA) remaining in beads and the length of attached fragments (17,26). *Eco*RI and *Hae*III cut cellular DNA in encapsulated cells into fragments that are multiples of 200 bp (Fig. 1I, lane 1). These fragments have a weight-average molecular weight of 3.1 kb, the length expected if the enzymes cut only in linker DNA (19). After removing most chromatin, the residual fragments are longer and the nucleosomal repeat is less obvious (Fig. 1I, lane 2); this is consistent with ~1 kb at attachment points being protected from cutting (19). A number-average molecular weight of 3.84 kb can be derived from the distribution of fragments in lane 2 (see Materials and Methods) and—as 4.3% chromatin remained attached—the average contour length is then $(100/4.3) \times 3.84 = 89.3$ kb. After trimming with eight other restriction enzymes, released fragments are shorter, but retain an obvious nucleosomal repeat of 182 bp (Fig. 1I, lane 3). The attached fragments are also shorter, but are still ~1 kb longer than the released ones, with a less obvious repeat (Fig. 1I, lane 4). The contour length also remains essentially unchanged, as the reduction in length is offset by increased detachment.

These results show that throughout the procedure the contour length is close to the average value of 86 kb obtained previously (17,19); therefore few attachments are made or broken during isolation. Note that nucleosomes cannot 'slide' along DNA because restriction sites that were initially covered remain covered—and so uncut—during the lengthy incubations. Note also that as attached fragments lack the obvious nucleosomal repeat typical of detached fragments, they must be relatively free of contamination by them.

Clones, 140, were then selected and inserts in 106 sequenced from each end. (See Materials and Methods for a discussion of the selection criteria.) Sixteen were eliminated as they were identical to others among the 106; they probably arose from repeated ligation of identical sequences in the original isolate, or as bacteria divided during amplification of the library. The resulting LASs were analysed for homology with sequences in the

databases (Table 1). Some proved to be mitochondrial, *alu*, alphoid or LINE repeats. Only two of the remainder were identified: LAS77 (part of the gene for the low-affinity Fc receptor; 27) and LAS95 (the non-transcribed region 5′ to the rDNA locus; 28). The latter region has been shown to be attached previously using this approach (24). Ten control inserts were also derived by fragmenting 'total' DNA with *Hae*III and then cloning the fragments.

Table 1. Properties of 90 LASs and 10 control clones containing 'total' DNA

	No. independent clones (%)	No. attached (%)[a]	No. transcribed (%)[b]
LASs			
Mitochondrial	25 (28)	3/3[c] (100)	3/3[c] (100)
Nuclear			
Unique	37 (41)	13/24[d] (54)	22/29[d] (76)
Alu	20 (22)	nt[e]	nt[e]
Alphoid	3 (3)	nt[f]	nt[f]
LINES	3 (3)	nt	nt
Other repeats	2 (2)	nt	nt
Controls	10 (100)	0/9[d] (0)	1/9[g] (11)

nt: not tested.

[a]Determined by 'Southern' blotting (as in Fig. 4); enrichments ≥2× are scored as 'attached'.

[b]Determined by 'Northern' blotting (as in Fig. 5).

[c]Three representative examples tested.

[d]Some sequences excluded because 'Southerns' gave no signals (perhaps because their targets were too short to blot efficiently), or 'Northerns' gave smears.

[e]One representative example tested; repeats not enriched and 'Northern' gave strong, repeated, signal.

[f]Most bands given by representative example, LAS14, not enriched or transcribed (Figs 4 and 5).

[g]One sequence gave repeats on 'Southern' that were not enriched; it was not used for 'Northern'.

Mitochondrial attachments

Mitochondrial clones accounted for 28% (Table 1), even though only ~0.15% DNA in a HeLa cell is mitochondrial (29). Clones tended to be derived from one of three regions of the genome (Fig. 2). This probably results, in part, from an appropriate distribution of *Alu*I and *Hae*III restriction sites within the three regions. These two enzymes cut efficiently under our conditions to generate 54 and 36%, respectively, of all ends cloned by our procedure (not shown) and, as the resulting mitochondrial fragments have the appropriate size, they are cloned preferentially and so over-represented in the library. But despite this efficient cutting, the three regions were nevertheless retained.

The same part of the 16S rRNA gene was present in four clones (i.e. C6, C8, C34 and C66), three of which must have arisen independently as the sequence was present in different lengths or orientations. Two other clones (i.e. C45 and C109) contained another region cloned in opposite orientations. Clearly, these two regions resisted elution more frequently than others, and inspection showed they shared the motif TAAG(N)$_5$GTGGGTTT. All other mitochondrial clones contained matches to the part-motif GTGGGTTT in at least six out of eight positions; however, loose matches to other 8mers were also as common.

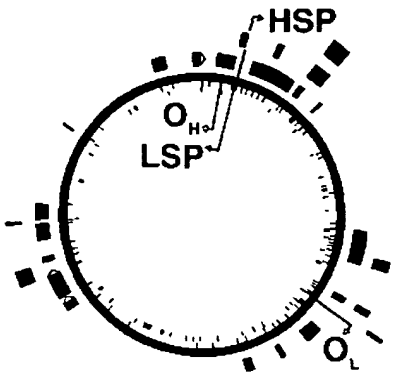

Figure 2. Positions of the mitochondrial clones are drawn outside the circular map of the genome; the heavy- and light-strand origins (O_H and O_L) and promoters (HSP and LSP), and the 64 *Alu*I and 50 *Hae*III sites (outer and inner ring of ticks) are also shown.

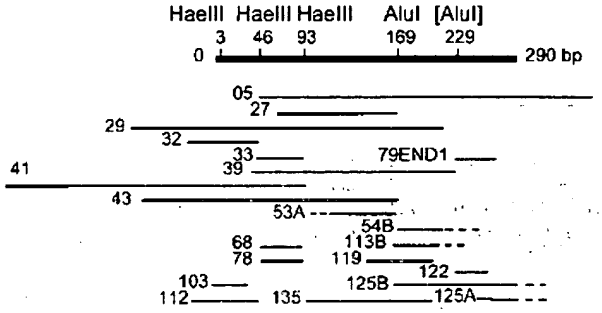

Figure 3. Maps of clones with homology with *alu*-a. The upper (thick) line represents the 290 bp *alu*-a sequence showing the positions of *Hae*III and *Alu*I sites (the site in square brackets is present in a variant, but not in *alu*-a). Overlapping lines below illustrate regions of homology with *alu*-a in the various LASs indicated (dotted lines: end of insert not sequenced).

Clones with repeated nuclear sequences

Although *alu* repeats represent ~2.5–5% of the nuclear genome, 31% non-mitochondrial clones contained *alu* repeats (Table 1). Despite this preferential attachment, no one region of the *alu* consensus sequence was invariably present (Fig. 3). Most repeats were of sub-family a (30,31) but members of sub-families b (e.g. LAS05, 119) and perhaps c (LAS33) were also present; none were 'precise variants' (32). LAS23 was almost identical to a repeated unit downstream of an *alu* repeat consisting of eight tandem repeats each with an *Alu*I site at the same location (33). Three inserts homologous with alphoid DNA shared positions 76–121 in the consensus sequence (Table 1; 34), perhaps suggesting a common attachment point, but this region did not contain the centromeric CENP-B box (positions 127–143; 35).

Confirmation that clones contained attached sequences

We next confirmed that the library did indeed contain attached sequences. Sequences remote from attachment points should be detached and so eluted from beads, unlike those close to

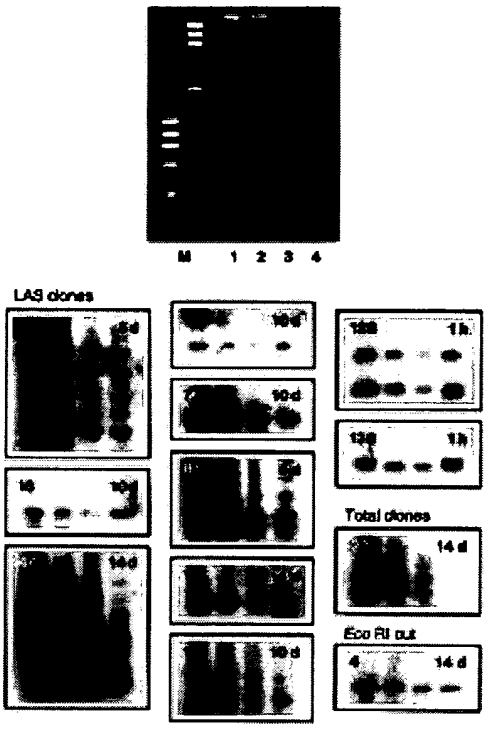

Figure 4. Mapping the proximity of different sequences to attachment points. Samples of total and attached DNA (prepared as in Fig. 1A–E) were applied to a number of gels. Lane M: φX174/*Hae*III and λ/*Hin*dIII markers. Lanes 1–3: 9, 3 and 1 μg total DNA. Lane 4: 1 μg attached fragments (9–12% total). After electrophoresis and staining with ethidium, gels were photographed under uv-illumination and a typical one is shown at the top. DNA in the other gels was blotted, hybridized with ^{32}P-labelled LASs or fragments of total DNA (clones 22 and 4; bottom right), and autoradiograms prepared. Clone number and exposure time are indicated in the top left and right of each panel. *Eco*RI cut: fragments prepared using only *Eco*RI (and not *Hae*III + *Eco*RI).

attachment points (5,36). Therefore we determined the relative concentrations of each cloned sequence in 'total' and 'attached' DNA fractions by 'Southern' blotting. As exhaustive cutting with 10 enzymes is so expensive, the large quantities of 'attached' fraction needed were prepared using lower concentrations of *Eco*RI and *Hae*III; as a result, it contained only ~10% of the total. Equal weights of the two fractions were completely cut with *Hae*III, resolved into discrete fragments by electrophoresis (e.g. Fig. 4 top, lanes 3,4), blotted on to a filter and probed with [^{32}P]DNA from each clone in turn. Sequences close to attachment points will be enriched in the 'attached' fraction and so will yield bands of greater intensity in the resulting autoradiograms (Fig. 4 bottom). The degree of enrichment is determined by reference to known amounts of total DNA run in adjacent channels in the gel (Fig. 4 top, lanes 1–3).

The approach is exemplified by reference to LAS15 (Fig. 4, middle left). One band was seen in 'total' DNA when it was probed with [^{32}P]DNA from this clone (lane 1); therefore it probably contains a unique sequence. The intensity of this band decreased as less DNA was applied to the gel (lanes 1–3). However, an equal weight of 'attached' DNA yielded a band of greater intensity (compare lanes 3 and 4), showing that this band was enriched in this fraction. The band intensity in lane 4 was intermediate between that seen in lanes 1 and 2 (which contain 9× and 3× the weight of DNA respectively), showing it was enriched 4–5×, confirming that the insert in this clone is indeed attached. Note, however, that if a sequence was permanently attached in all cells it should be enriched 10× in an 'attached' fraction representing 10% of the total. However, no sequence analysed was enriched to this extent, suggesting that all were attached in only a fraction of the cells.

LAS73 and LAS77 (encoding the Fc receptor) were typical of many unique sequences, being enriched ~2× (Fig. 4). LAS117 provides another example of a highly-enriched (unique) sequence; as we shall see, its transcripts are found in high concentration. LAS124 detects two sequences, but only one is enriched. Repeated nuclear sequences give multiple bands of higher intensity after shorter exposures; for example, LAS14 (an alphoid repeat) gives six bands, five of which are not enriched. All of the many bands given by LAS40 (an unknown repeat) are depleted, as is the 'smear' given by LAS81, despite an enrichment of specific bands. These results suggest that some repeats in a family lie closer to attachment points than others. Mitochondrial sequences (exemplified in Fig. 4 by LAS8 and 62) give intense bands after short exposures; all were highly enriched (LAS8 and 62 were enriched 7.5 and 4×, respectively). *Alu* repeats also gave intense signals after short exposures but, in contrast, were not enriched (not shown); therefore it seems that only a sub-fraction are attached.

We found that 54% inserts with unique sequences were enriched in the 'attached' fraction (Table 1), confirming that they truly were attached. (Note sequences are conservatively scored as 'attached' only if they gave enrichments of 2-fold or more.) In contrast, none of the inserts in control clones were enriched (Fig. 4, 'total' clones 22 and 4; Table 1).

Many LASs are transcribed

The concentration of transcripts that were complementary to different LASs was determined by 'Northern' blotting. RNA from whole cells, nuclei or the polyA$^+$ fraction was separated by electrophoresis (Fig. 5 left, lanes 1–3, respectively), blotted and hybridized with ^{32}P-labelled probes prepared from clones containing either LASs or 'total' DNA; autoradiograms involving the same examples used in Figure 4 are illustrated in Figure 5. Two repeats (i.e. LAS14 and 40) were not transcribed, but another (LAS81) gave a strong signal with polyA$^+$ RNA and was the only LAS to hybridize with nuclear RNA. Of the single-copy sequences, LAS15 gave two weak bands in the polyA$^+$ fraction (arrows), but LAS73 was typical of most, giving a band in whole cell RNA and a stronger band in the polyA$^+$ fraction. LAS77 (the Fc receptor) hybridized with two faint bands in the polyA$^+$ fraction (arrows). [cDNAs encoding the Fc receptor have been isolated from many other libraries made from various tissues, showing that it is widely expressed (20).] LAS117 gave strong bands and 124 was typical of several LASs (including 35, 93, 101, 124 and 130) that hybridized with bands of different sizes in whole cell and polyA$^+$ RNA. The mitochondrial LASs, 8 and 62, are highly transcribed and give strong signals after short exposures; transcripts from the latter, but not the former, are poly-adenylated (as expected). (The lack of signal given by LAS8 with the polyA$^+$ and nuclear fractions reflects their purity.) None of the 'total' clones tested (e.g. 22) gave significant signals, even

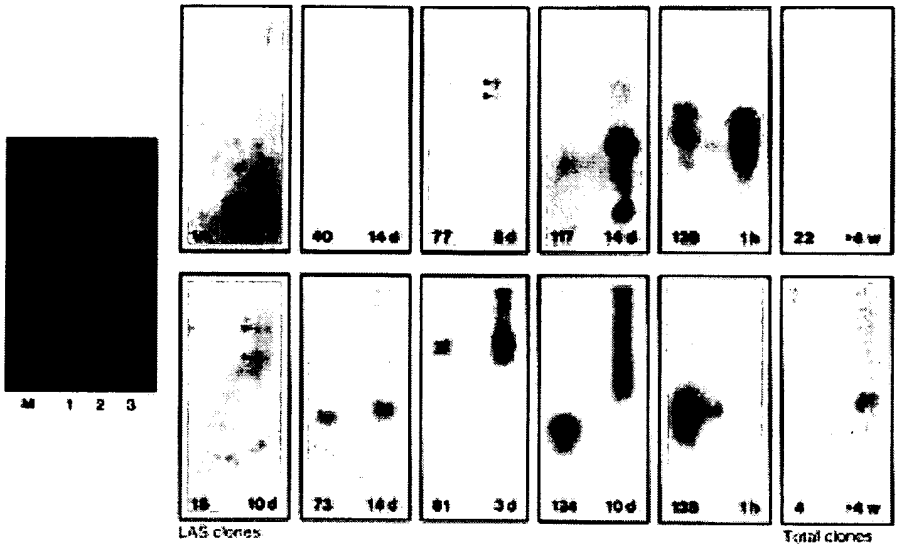

Figure 5. Most LASs are transcribed. Marker DNA, total cell RNA (25 µg), nuclear RNA (10 µg) and polyA$^+$ RNA (10 µg) were applied to lanes M, and 1–3, of a number of denaturing gels. (Relative cell equivalents of RNA loaded are roughly 1:5:12, respectively.) After electrophoresis and staining with ethidium, gels were photographed under uv-illumination and a typical one is shown on the left; the two rRNA bands (lanes 1,2) are not found in polyA$^+$ RNA (lane 3). RNA in the other gels was blotted on to nitrocellulose, hybridized with ^{32}P-labelled LASs or fragments of total DNA (clones 22 and 4; far right), and autoradiograms prepared. Clone number and exposure time are indicated in the bottom left and right of each panel.

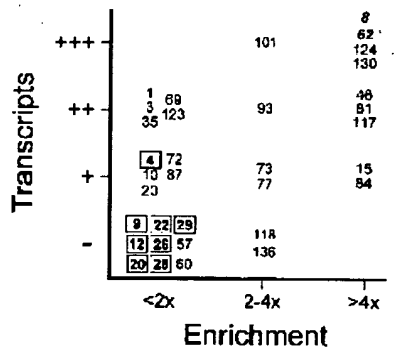

Figure 6. The relationship between the degree of enrichment (i.e. attachment) of the unique LASs relative to their transcript concentrations. Enrichments and transcript concentrations, ranging from non-detectable (–) to high (+++), were determined as in Figures 4 and 5. Results for mitochondrial LASs and clones containing 'total' DNA are also shown in bold in italics and boxes, respectively.

after long exposures using a sensitive PhosphorImager; 'total' clone 4 gave the strongest, but this was hardly above background.

The results of this analysis are summarized in Table 1. Figure 6 illustrates the rough correlation between the enrichment of both mitochondrial and nuclear LASs and the concentration of their transcripts; control fragments (shown in boxes) tend not to be enriched or transcribed.

Possible attachment motifs in nuclear LASs

We next screened the nuclear LASs containing unique sequences for possible attachment motifs. The same sequences 'shuffled' at random provided suitable controls with the same base content. No LAS contained an exact match with the (loosely-defined) topoisomerase II consensus sequence (37) and only seven contained the same sequence with ≤2 mismatches. [These values can be compared with 0 (exact match) and 9 (two mismatches) given by the 'shuffled' set.] One (none of the 'shuffled' set) had ≤2 mismatches with the topoisomerase I consensus sequence (38). Eighteen (19 'shuffled') had ≤2 mismatches with the SAR A box (39), whilst 15 (15 'shuffled') had ≤1 mismatch with the SAR T box; 12 (11 'shuffled') shared these loose matches with both boxes. Similarly, LASs and their 'shuffled' counterparts contained the same (few) number of SATB1 binding motifs (12). These results show LASs contain no more of these AT-rich motifs than are expected by chance. LASs were also screened (see Materials and Methods) for (i) deviations in base composition, hairpin loops, repeats, Z DNA, and common 'words'; (ii) promoters; and (iii) the binding sites for the transcription factors Sp1, AP2 and OCT1 that are known to be present in HeLa cells (40); no significant differences were found between the LASs and their 'shuffled' counterparts.

DISCUSSION

The approach

We have isolated and cloned DNA sequences that attach nuclear and mitochondrial DNA in loops to underlying structures. HeLa cells were encapsulated in agarose beads to protect them during subsequent manipulation and permeabilized with Triton X-100 in a 'physiological' buffer; then loops were cut with restriction enzymes, most DNA removed by electrophoresis and the residual 1.4% that remained attached was isolated and cloned, and 106 representative clones sequenced (Fig. 1). It is important to note

that these LASs are defined operationally by the procedure used; any attachments sensitive to Triton or the electric field would not be seen.

Chromatin is notoriously prone to rearrangement, and our approach requires that few existing attachments are broken nor new attachments created during the procedure. Therefore, we used isotonic conditions throughout. We have previously shown that the chromatin fibre in these permeabilized cells remains intact and that nuclear RNA and DNA synthesis can continue at roughly the rates found *in vivo* (16); if chromatin aggregated artifactually on lysis we would expect to lose activity. We also monitored the length of the nucleosomal repeat as well as the contour length of the nuclear loops; both remained unchanged during isolation. Moreover, nucleosomes did not 'slide' along DNA under these conditions because restriction sites that were initially covered remain covered, and so uncut, during lengthy incubations. And as the attached fragments lacked the obvious nucleosomal repeat typical of detached fragments, they must have been relatively free of contamination by them. We also confirmed that we had indeed cloned attached sequences by measuring the relative concentrations of complementary sequences in 'attached' and 'total' DNA fractions (Fig. 4); 54% of the clones tested but none of the controls behaved in the expected manner (i.e. homologous sequences were enriched in the 'attached' fraction; Table 1).

Attachments of mitochondrial DNA

Mitochondria are associated with cytoskeletal elements that resist elution from beads (25) and so we expected to clone mitochondrial sequences. However, 28% of our clones were derived from this organelle (Table 1), even though only ~0.15% of the DNA in the cell is mitochondrial (29). Even more surprisingly, clones tended to be derived from one of three functionally important regions (Fig. 2). One region includes the origin of replication of the heavy strand, the heavy- and light-strand promoters and the highly-transcribed rRNA genes, the second the other origin, and the third a region that is commonly mutated or deleted in the mitochondrial myopathies (41).

'Structural' and 'functional' attachments

As discussed in the Introduction, two extreme kinds of attachment can be imagined, 'structural' and 'functional', and we expected to find some combination of the two. *Alu* sequences, which comprise 2.5–5% of the genome, constituted 31% of the non-mitochondrial LASs (Table 1) and so become candidates for members of the 'structural' class. However, Southern blotting showed that most *alu* repeats were not preferentially associated with the substructure (not shown); therefore only a fraction, presumably represented by the ones cloned, can be involved in attachment. Moreover, no one region within the repeat was invariably present in the *alu* LASs (Fig. 3), so it seems unlikely that they contain a 'structural' motif.

A 'structural' motif that was permanently bound in all cells in the population should be enriched 10-fold when all but 10% chromatin was detached. However, none of our nuclear or mitochondrial LASs were enriched to this extent (Fig. 6). The enrichments, which are generally 1.5–5×, imply that most LASs were attached in only a fraction (i.e. 15–50%) of loops in the population; they only have a higher probability than others of being attached. By definition, then, they are 'functional'.

Although our LASs are characterized by their diversity, they do share one common feature: most are transcription units and so again can be classified as 'functional'. For example, 76% unique LASs were transcribed compared to ~5% human genome (42) and only one of nine controls (Table 1). In addition, all regions of the mitochondrial genome, and so all mitochondrial LASs, are also transcribed (43). [This was confirmed for only two (Fig. 6, LAS8 and 62).] Moreover, it is also possible, but difficult to prove, that most *alu* LASs are transcriptionally active, since it is now known that many *alu* repeats are transcribed; they occur within other transcription units and some are even functional polymerase III genes with internal promoters and operative retinoic acid response elements (e.g. 44–46). Even if the attached *alu*s are not transcribed, they nevertheless carry the stigmata of a transcription unit. Strikingly, then, most of our LASs—whether they be unique, mitochondrial or *alu* sequences—turn out to be parts of transcription units.

Replicating sequences, and perhaps origins, are closely associated with nuclear matrices and cages (1,2,50). Although <1% loops in our unsynchronized cells can be replicating at any one time, it remains formally possible that origins mediate many attachments. However, it is difficult to test this in the absence of appropriate assays.

Attachments seen in different material

There is an extensive literature on three kinds of nuclear substructure isolated using hypo- and hyper-tonic salt concentrations, scaffolds, matrices and cages. Each is different from the others and each is associated with a different subset of genomic DNA (for reviews, see 2,15). Our attachment sequences are unlike those associated with scaffolds (3); they contain no more topoisomerase II sites or SAR consensus sequences than would be expected from their base composition. SARs and MARs are also very AT rich, but our non-mitochondrial LASs were less AT rich (i.e. 54%) than the whole genome (i.e. 60%). MARs are sometimes transcribed (47), sometimes not (48), depending on the precise procedure used for isolation; most of our LASs are transcribed (Table 1). Unlike MARs (12,49), our LASs contain no more inverted repeats or SATB1-binding sequences than 'shuffled' controls. This lack of similarity with either SARs or MARs is unsurprising as loops decrease in length during the preparation of scaffolds and matrices (17); this means that the new attachments that create the smaller loops will inevitably obscure pre-existing ones.

The attachments seen here, however, are similar to those seen earlier in cages prepared by lysing HeLa cells in Triton and 2 M NaCl without prior isolation of nuclei in hypotonic buffers. The resulting nucleoids contained loops of naked, supercoiled, DNA attached to the cage; again mitochondrial DNA was preferentially attached, no one repeated sequence was responsible for attaching the nuclear loops, and attached unique sequences tended to be transcribed or were rich in promoter or enhancer sequences (50). This similarity was expected as loop length increases slightly on isolation, so it is unlikely that new, obscuring, attachments were created (17).

The attachments of a (nuclear) 'minichromosome' containing two transcription units have also been analysed using isotonic conditions and the approach illustrated in Figure 1 (51). Two

populations of minichromosomes were found; one was transcriptionally inactive and could be eluted, the other was active but resisted elution. The resistant minichromosomes were attached, on average, at only one point, either through a promoter or a transcription unit; this suggested that they were bound either through transcription factors (at a promoter) or an engaged polymerase (on a transcription unit). Cutting attached minichromosomes with *Hae*III enabled most resulting fragments to elute without loss of polymerizing activity; then the residual fragments still served as templates for engaged polymerases. This suggested that minichromosomal attachments changed from moment to moment as the template engaged an attached polymerase.

'Functional' attachments to transcription 'factories'

The simplest explanation for the results obtained here, and those obtained earlier with cages and minichromosomes, is as follows. We now know that active RNA polymerases are concentrated in HeLa nuclei in ~2100 discrete structures (diameter ~70 nm) attached to a nucleoskeleton (52,53; F. J. Iborra, A. Pombo, D. A. Jackson and P. R. Cook, manuscript submitted). Each of these 'factories' is associated with many loops and transcription units. Initiation would involve attachment of a promoter first to transcription factors and then a polymerase in a factory, before the transcription unit slid through the polymerization site and nascent RNA was extruded into the factory; after the completion of the transcript, the unit would detach from the factory. At any moment individual templates would be at different stages in this cycle; no one point would always be attached and different points would have different probabilities of attachment, depending on promoter strength and polymerase transit time. Enhancers (and any non-transcribed *alu* repeats) would also have a high probability of binding to a factory, but would do so without initiating transcription. All such attachments are dynamic in the sense that they continually change from moment to moment (54,55).

A similar kind of model can be extended to mitochondria. The nucleolar and mitochondrial transcription systems share a common ancestry (56) and as nucleoli also contain transcription 'factories' located on the nucleoskeleton (57), it seems likely that mitochondria will contain homologous structures attached to the cytoskeleton. Then the transcriptionally-active fraction of mitochondrial genomes would also be attached through promoters or transcription units to the underlying structure.

ACKNOWLEDGEMENTS

We thank Sandra Smith and Liz Cowe for their help, and the Cancer Research Campaign for support.

REFERENCES

1 Getzenberg,R.H., Pienta,K.J., Ward,W.S. and Coffey,D.S. (1991) *J. Cell. Biochem.*, **47**, 289–299.
2 Jack,R.S. and Eggert,H. (1992) *Eur. J. Biochem.*, **209**, 503–509.
3 Roberge,M. and Gasser,S.M. (1992) *Mol. Microbiol.*, **6**, 419–423.
4 Jackson,D.A., Dolle,A., Robertson,G. and Cook,P.R. (1992) *Cell Biol. Int. Rep.*, **16**, 687–696.
5 Cook,P.R. and Brazell,I.A. (1980) *Nucleic Acids Res.*, **8**, 2895–2906.
6 Mirkovitch,J., Mirault,M.-E. and Laemmli,U.K. (1984) *Cell*, **39**, 223–232.
7 Cockerill,P.N. and Garrard,W.T. (1986) *Cell*, **44**, 273–282.
8 Grosveld,F., van Assendelft,G.B., Greaves,D.R. and Kollias,G. (1987) *Cell*, **51**, 975–985.
9 Stief,A., Winter,D.M., Stratling,W.H. and Sippel,A.E. (1989) *Nature*, **341**, 343–345.
10 Kellum,R. and Schedl,P. (1991) *Cell*, **64**, 941–950.
11 Adachi,Y., Käs,E. and Laemmli,U.K. (1989) *EMBO J.*, **8**, 3997–4006.
12 Dickinson,L.A., Joh,T., Kohwi,Y. and Kohwi-Shegematsu,T. (1992) *Cell*, **70**, 631–645.
13 Romig,H., Fackelmayer,F.O., Renz,A., Ramsperger,U. and Richter,A. (1992) *EMBO J.*, **11**, 3431–3440.
14 von Kries,J.P., Buhrmester,H. and Stratling,W.H. (1991) *Cell*, **64**, 123–135.
15 Cook,P.R. (1988) *J. Cell Sci.*, **90**, 1–6.
16 Jackson,D.A., Yuan,J. and Cook,P.R. (1988) *J. Cell Sci.*, **90**, 365–378.
17 Jackson,D.A., Dickinson,P. and Cook,P.R. (1990) *EMBO J.*, **9**, 567–571.
18 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
19 Jackson,D.A., Dickinson,P. and Cook,P.R. (1990) *Nucleic Acids Res.*, **18**, 4385–4393.
20 Adams,M.D. *et al.* (1995) *Nature*, **377** supplement, 3–17.
21 Staden,R. (1989) *CABIOS*, **5**, 293–298.
22 Staden,R. (1990) *Methods Enzymol.*, **183**, 193–211.
23 Prestridge,D.S. (1995) *J. Mol. Biol.*, **249**, 923–932.
24 Dickinson,P., Cook,P.R. and Jackson,D.A. (1990) *EMBO J.*, **9**, 2207–2214.
25 Hozák,P., Jackson,D.A. and Cook,P.R. (1994) *J. Cell Sci.*, **107**, 2191–2202.
26 Igo-Kemenes,T. and Zachau,H.G. (1977) *Cold Spring Harbor Symp. Quant. Biol.*, **42**, 109–118.
27 Brooks,D.G., Qiu,W.Q., Luster,A.D. and Ravetch,J.V. (1989) *J. Exp. Med.*, **170**, 1369–1385.
28 Sylvester,J.E., Petersen,R. and Schmickel,R.D. (1989) *Gene*, **84**, 193–196.
29 Nass,M.M.K. (1969) *J. Mol. Biol.*, **42**, 521–528.
30 Jurka,J. and Smith,T. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 4775–4778.
31 Batzer,M.A., Kilroy,G.E., Richard,P.E., Shaikh,T.H., Deselle,T.D., Hoppens,C.L. and Deininger,P.L. (1990) *Nucleic Acids Res.*, **18**, 6793–6798.
32 Matera,A.G., Hellman,U. and Schmid,C.W. (1990) *Mol. Cell. Biol.*, **10**, 5424–5432.
33 Yang,R., Fristensky,B., Deutch,A.H., Huang,R.C.C., Tan,Y.H., Narang,S.A. and Wu,R. (1983) *Gene*, **25**, 59–66.
34 Choo,K.H., Vissel,B., Nagy,A., Earle,E. and Kalitsis,P. (1991) *Nucleic Acids Res.*, **19**, 1179–1182.
35 Muro,Y., Masumoto,K., Yoda,K., Nozaki,N., Ohashi,M. and Okazaki,T. (1992) *J. Cell Biol.*, **116**, 585–596.
36 Cook,P.R. Lang,J., Hayday,A., Lania,L., Fried,M., Chiswell,D.J. and Wyke,J.A. (1982) *EMBO J.*, **1**, 447–452.
37 Sander,M. and Hsieh,T.S. (1985) *Nucleic Acids Res.*, **13**, 1057–1072.
38 Bonven,B.J., Gocke,E. and Westergaard,O. (1985) *Cell*, **41**, 541–551.
39 Gasser,S.M., Amati,B.B., Cardenas,M.E. and Hoffman,J.F.-X. (1989) *Int. Rev. Cytol.*, **119**, 57–96.
40 Mitchell,P.J. and Tjian,R. (1989) *Science*, **245**, 371–378.
41 Wallace,D.C. (1992) *Annu. Rev. Biochem.*, **61**, 1175–1212.
42 Lewin,B. (1975) *Cell*, **4**, 77–93.
43 Clayton,D.A. (1991) *Annu. Rev. Cell Biol.*, **7**, 453–478.
44 Maraia,R.J., Driscoll,C.T., Bilyeu,T., Hsu,K. and Darlington,G.J. (1993) *Mol. Cell. Biol.*, **13**, 4233–4241.
45 Panning,B. and Smiley,J.R. (1995) *J. Mol. Biol.*, **248**, 513–524.
46 Vansant,G. and Reynolds,W.F. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 8229–8233.
47 Basler,J., Hastie,N.D., Pietras,D., Matsui,S.I., Sandberg,A.A. and Berezney,R. (1981) *Biochemistry*, **20**, 6921–6929.
48 Ciejek,E.M., Tsai,M.J. and O'Malley,B.W. (1983) *Nature*, **306**, 607–609.
49 Boulikas,T. and Kong,C.F. (1993) *J. Cell. Biochem.*, **53**, 1–12.
50 Jackson,D.A., McCready,S.J. and Cook,P.R. (1984) *J. Cell Sci. Suppl.*, **1**, 59–79.
51 Jackson,D.A. and Cook,P.R. (1993) *J. Cell Sci.*, **105**, 1143–1150.
52 Jackson,D.A., Hassan,A.B., Errington,R.J. and Cook,P.R. (1993) *EMBO J.*, **12**, 1059–1065.
53 Wansink,D.G., Schul,W., van der Kraan,I., van Steensel,B., van Driel,R. and de Jong,L. (1993) *J. Cell Biol.*, **122**, 283–293.
54 Cook,P.R. (1994) *BioEssays*, **16**, 425–430.
55 Cook,P.R. (1995) *J. Cell Sci.*, **108**, 2927–2935.
56 Parisi,M.A. and Clayton,D.A. (1991) *Science*, **252**, 965–969.
57 Hozák,P., Cook,P.R., Schöfer,C., Mosgöller,W. and Wachtler,F. (1994) *J. Cell Sci.*, **107**, 639–648.